



# UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
[www.uspto.gov](http://www.uspto.gov)

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/905,610	07/13/2001	Maria Castellanos	10007911	3840

7590 04/29/2005  
HEWLETT-PACKARD COMPANY  
Intellectual Property Administration  
P.O. Box 272400  
Fort Collins, CO 80527-2400

EXAMINER

LUDWIG, MATTHEW J

ART UNIT	PAPER NUMBER
----------	--------------

2178

DATE MAILED: 04/29/2005

Please find below and/or attached an Office communication concerning this application or proceeding.

# Office Action Summary

Application No.

09/905,610

Applicant(s)

CASTELLANOS ET AL.

Examiner

Matthew J. Ludwig

Art Unit

2178

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

## Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

## Status

- 1) ☒ Responsive to communication(s) filed on 13 July 2001.
- 2a) ☐ This action is **FINAL**. 2b) ☒ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

## Disposition of Claims

- 4) ☒ Claim(s) 1-24 is/are pending in the application.
- 4a) Of the above claim(s) \_\_\_\_\_ is/are withdrawn from consideration.
- 5) ☐ Claim(s) \_\_\_\_\_ is/are allowed.
- 6) ☒ Claim(s) 1-5, 9-13 and 17-21 is/are rejected.
- 7) ☒ Claim(s) 6-8, 14-16 and 22-24 is/are objected to.
- 8) ☐ Claim(s) \_\_\_\_\_ are subject to restriction and/or election requirement.

## Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☐ The drawing(s) filed on \_\_\_\_\_ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.  
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).  
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

## Priority under 35 U.S.C. § 119

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All b) ☐ Some \* c) ☐ None of:
- ☐ Certified copies of the priority documents have been received.
  - ☐ Certified copies of the priority documents have been received in Application No. \_\_\_\_\_.
  - ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

\* See the attached detailed Office action for a list of the certified copies not received.

## Attachment(s)

- ☒ Notice of References Cited (PTO-892)
- ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
- ☐ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)  
Paper No(s)/Mail Date \_\_\_\_\_
- ☐ Interview Summary (PTO-413)  
Paper No(s)/Mail Date. \_\_\_\_\_
- ☐ Notice of Informal Patent Application (PTO-152)
- ☐ Other: \_\_\_\_\_

### DETAILED ACTION

1. This action is responsive to communications: Application filed 8/17/01.
2. Claims 1-24 are pending in the application and claims 1, 9, and 17, are independent claims.

#### *Claim Rejections - 35 USC § 103*

3. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

4. **Claims 1-5, 9-13, and 17-21, are rejected under 35 U.S.C. 103(a) as being unpatentable over Thompson et al., Pub. No. US 2002/0103834 filed (6/27/01).**

**In reference to independent claim 1, Thompson teaches:**

Analyzing electronic forms in several stages including cleaning text images, error correction of ASCII text, and data mining of the cleaned text. See Thompson, figure 1. The reference provides the suggestion of locating valid and invalid terms, which are words within a document. The terms could include phrases that produce reference words and non-reference words from a selected group of documents (compare to “*generating a list of reference words and phrases and a list of non-reference words and phrases from a selected group of documents*”). See page 16, [0538] through [0546].

The solution and correction method taught by Thompson provides a similar invalid character correction algorithm to produce possible solutions to invalid terms found within the

Art Unit: 2178

document (compare to “*comparing said list of reference words and phrases with a joined list containing said reference...*”). See page 16, [0544] through [0546].

The system’s error correction performance can be improved by selecting reference dictionaries that are more complete relative to the lexicon of the working documents. Users can do this by selecting/unselecting reference dictionaries.

The system allows the user to locate these terms and to build “user generated” dictionaries. Users can build “user generated” dictionaries by selecting and saving terms out of the working documents, all-alpha exceptions list (compare to “*filtering said approximate duplicates list to create a thesaurus of standard words*”). See page 15, [0521] through [0532].

Allows users to convert standard ASCII text and database files into the format of the system’s reference databases. Users can use this utility to convert lists of terms, free texts, and standard computer databases into system accessible reference databases (compare to “*editing said selected group of documents with an editor operable to use said thesaurus to replace a word or phrase on said approximate duplicates...*”). See page 16, [0536] through [0546].

The correction methods taught by Thompson do not explicitly point to the teaching of a Thesaurus, however, the suggestion of a system containing four resident reference dictionaries which users may select or un-select, one of which specifically mentions commonly used words, provides a proficient description of a similarly recited Thesaurus. Both locate words within a document and provide a user with a similar word based on disparate matches. Therefore it would have been obvious to one of ordinary skill in the art at the time the invention was made to have an external reference dictionary and generated a Thesaurus from the dictionary containing commonly used words because it would have given the user enhanced system performance by

removing non-validated “correct” terms from the exceptions lists and allows the system to generate more accurate statistical profiles of the set of working documents.

**In reference to dependent claim 2, Thompson teaches:**

The user instructs the system not to process documents in which more than 25% of the terms are invalid. The error resolution system further distinguishes invalid terms (i.e. non-words) as exceptions meaning terms with a majority of alpha characters, and trash. See page 14, [0492] through [0494].

**In reference to dependent claim 3, Thompson teaches:**

Having converted the user-selected working documents into indexes, the system compares terms in these indexes with the terms in the user-selected reference dictionaries. Terms that are left after these two screening processes are either exceptions or trash. See page 14, [0483] through [0488].

**In reference to dependent claim 4, Thompson teaches:**

Analyzing electronic forms in several stages including cleaning text images, error correction of ASCII text, and data mining of the cleaned text. See Thompson, figure 1. The reference provides the suggestion of locating valid and invalid terms, which are words within a document. The terms could include phrases that produce reference words and non-reference words from a selected group of documents. See page 16, [0538] through [0546].

Document quality ratings impact solution confidence values. Documents with high quality ratings have fewer invalid terms, fewer non-validated terms, and more complete sets of valid terms. For documents with these characteristics, the confidence values of the possible solutions located by the system are accepted at par. By the same token, documents with lower

quality ratings have more invalid terms, more non-valid terms, and less complete sets of valid terms. See page 18, [0620] through [0623].

**In reference to dependent claim 5**, Thompson teaches:

Because working documents contain terms that are not in the system's resident dictionaries the system allows the user to locate these terms and to build user-generated dictionaries. See page 15, [0526] through [0533].

**In reference to claims 9-13**, the limitations reflect the system comprising instructions used for performing the methods as claimed in 1-5, respectively, and in further view of the following, are rejected along the same rationale.

**In reference to claims 17-21**, the limitations reflect the system comprising computer readable instructions used for performing the methods as claimed in 1-5, respectively, and in further view of the following, are rejected along the same rationale.

### ***Allowable Subject Matter***

Claims 6-8, 14-16, and 22-24, are objected to as being dependent upon a rejected base claim, but would be allowable if rewritten in independent form including all of the limitations of the base claim and any intervening claims.

### ***Conclusion***

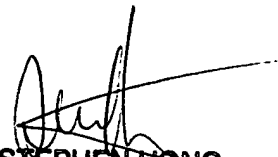
Any inquiry concerning this communication or earlier communications from the examiner should be directed to Matthew J. Ludwig whose telephone number is 571-272-4127. The examiner can normally be reached on 8:00am-5:00pm.

Art Unit: 2178

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Stephen Hong can be reached on 571-272-4124. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

ML  
April 23, 2005



**STEPHEN HONG**  
SUPERVISORY PATENT EXAMINER